DOCUMENT RESUME

ED 134 927 CS 003 180

AUTHOR	Goetz, Ernest T.
TITLE	Sentences in Lists and in Connected Discourse. Technical Report No. 3.
INSTITUTION	Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	Nov 75
GRANT	NIE-G-74-0007
NOTE	75p.
EDRS PRICE	MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS	*Connected Discourse; Elementary Secondary Education; *Experiments; Higher Education; Learning Processes; Memory; Reading Comprehension; *Reading Research; *Recall (Psychological); *Sentences
IDENTIFIERS	*Center for the Study of Reading (Illinois)

ABSTRACT

The recent deluge of published studies employing sentences or connected discourse as the unit of study has left unsolved the question of whether the two types of materials are essentially similar or importantly different. An understanding of this issue is crucial to theory, since the existence of major psychological differences between the comprehension, learning, and memory of sentence lists and connected discourse would make a unified theory covering both types of materials exceedingly difficult to formulate. While offering no final resolution of the issue, the present paper examines the evidence, considers the implications of several major theoretical developments, and poses questions amenable to experimentation. It is hoped that the paper will serve as a springboard to a higher level of understanding of how people process these two common types of experimental materials. A list of references is included. (Author/JM)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED134927

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Technical Report No. 3

SENTENCES IN LISTS AND IN CONNECTED DISCOURSE

Ernest T. Goetz

November 1975

Laboratory for Cognitive Studies in Education

University of Illinois at Urbana-Champaign

The research reported herein was supported in part by the National Institute of Education under Contract HEW NIE-G-74-0007.

CS 003 180

Abstract

The recent deluge of published studies employing sentences or connected discourse as the unit of study has left the question of whether the two types of materials are essentially similar or importantly different unsolved. An understanding of this issue is crucial to theory, since the existence of major psychological differences between the comprehension, learning, and memory of sentence lists and connected discourse would make a unified theory covering both types of materials exceedingly difficult to formulate.

While offering no final resolution of the issue, the present paper examines the evidence, considers the implications of several major theoretical developments, and poses questions amenable to experimentation. It is hoped that the paper will serve as a springboard to a higher level of understanding of how people process these two common types of experimental materials.

Sentences in Lists and in Connected Discourse

The last few years have seen a marked trend in the literature of experimental psychology toward the study of the comprehension, learning, and memory of connected discourse. The recent surge of interest seems to have begun with papers by Fillenbaum (1966) and Sachs (1967), but the area is not exactly new to psychology. Ebbinghaus (1885) himself studied discourse learning and reported that he needed only one-tenth the number of repetitions to learn a passage of prose (from "Don Juan") as he did to learn a list of nonsense syllables equal in length to the number of syllables in the prose. Binet and Henri (1894) reported a series of studies on discourse memory with children which showed great insight and anticipated several of the major themes of recent experimental work.

Although the tradition of discourse experiments is far older, recent developments have caused the swift formation of a literature and tradition of experiments on sentences in unrelated lists. The event which marks the beginning of the heyday of sentence list experiments was the creation of the discipline of psycholinguistics. Early experiments with sentence lists (e.g., Mehler, 1963; Miller, 1962) and most since then have been concerned with attempting to test experimentally the psychological validity of linguistic concepts such as transformational complexity (Chomsky, 1957, 1965) or Yngve depth (1960).

While the study of sentences qua sentences is certainly justifiable on its own merits, the relationship between this newer experimental material and connected discourse remains unclear. Deese (1961) considered the even broader question of the relationship of experimental materials "from the isolated verbal unit to connected discourse," but, unfortunately,

his paper was written before the recent deluge of work and was influenced primarily by work done with approximations to English (Miller & Selfridge, 1950). It is, perhaps, time to reexamine the question in light of more recent theory and experiment.

The implicit assumption that seems to run through most of the work in the area is that both sentence lists and connected discourse are essentially similar psychologically, although discourse, being more complex, may involve additional processes. This sentiment is reflected in the post-Chomskian orthodox position that linguistic accounts of the sentence can be extended to discourse because sentences in discourse can be joined together by appropriate conjunctions or other connectives to form one extended hypersentence (e.g. Katz & Fodor, 1963).

While this view is not implausible, there is reason to question it. Bartlett (1932) warned that the apparent simplicity of experimental materials, when presented to so complex an organism as a human subject, does not ensure simplicity of psychological processes. This has recently proved to be the case with the traditional materials of verbal learning. The work of Montague, Adams, Prytullac and others (see Montague, 1972 for review) shows that such seemingly simple materials as nonsense syllables, paired-associates, and word lists are often processed and learned in highly complex and idiosyncratic fashion. Therefore, it is surely risky to simply assume that learning sentences in lists is similar to and simpler than learning connected discourse.

The purpose of this paper is to examine the question of whether the processes involved in learning sentences in lists and connected discourse

are truly similar or whether important differences exist. A better understanding of this relationship is crucial in order to permit the interpretation of the burgeoning body of literature in the sentence-discourse area. A better understanding would also help to guide decisions of future experimenters as to whether an experiment should employ a sentence list or discourse: if the two are essentially similar with regard to the psychological processes in question, then such a decision could be made on the basis of expedience or other considerations, but if there are important differences, then the decision will have to be made on the basis of an understanding of those differences. Further, such an understanding should guide the scope of future theories, whether they should cover both types of experimental materials or be specific to one or the other. The understanding gained might even influence the content of future theories by spotlighting important psychological aspects of the materials.

Since there is a paucity of experimental data on the issue, it will not be possible to achieve a resolution of the question within the confines of this paper. Rather, I will attempt to provide background for considering the question, by reviewing a few important similarities between sentence lists and discourse, as a backdrop against which to view the discussion of established and possible differences, which will constitute the bulk of the paper. This will be followed by a more detailed examination of the relatively few studies which have directly compared the two materials. After that, a more explicit statement of the questions involved will be made, and the issue will be examined in the light of theoretical and empirical advances in the fields of discourse and memory structure and subject

processing. Finally, specific hypotheses and suggestions for their experimental investigation will be offered. In this way, it is hoped that the paper will help move psychology some small distance further down the road to understanding what happens when people read or hear sentences and connected discourse.

Similarities Between Sentence Lists and Discourse

Memory for Ideas Versus Memory for Words

Experimental psychology, whether due to the influence of years of experimentation with nonsense syllables, for which paraphrase is impossible, or due to the procedural difficulties of gist scoring, has tended to study verbatim reproduction of sentences and discourse instead of memory for the meaning of the sentences. People can, to be sure, remember sentences and discourse verbatim. Nearly everyone can recite word for word a few proverbs, several nursery rhymes, and any number of slogans for products advertised on television. Actors can commit to memory the exact wording of whole scripts. People can remember not only the words, but incidental physical attributes of the form in which they were presented. Thus, for example, people can recognize the presentation mode (audio or visual) of words (e.g. Bray and Batchelder, 1972; Hintzman, Block, and Inskeep, 1972). Memory for the gender of the reader's voice has been demonstrated with words (e.g. Hintzman, et al., 1972) and sentences (Light, Strasburg, Rubin, and Linde, 1973). Hintzman, et al. showed the subjects' ability to discriminate between words presented in different type faces (upper case block letters versus lower case script) and Kolers and Ostry (1974) found an effect of typography (normal versus inverted) of the initial

presentation of sentences upon recognition, even after a 32 day retention interval. In the only study of this type using discourse, Rothkopf (1971) proved that subjects were able to specify, with greater than chance accuracy, the location of the page on which information from a passage occurred with respect to the entire 300 word passage, and also the spatial location of the information on the page.

Although people clearly can remember the exact words and even the physical form of the words, it is also unquestionably true and of paramount psychological importance that people usually remember the idea of what they hear or read far longer and much more easily than the words that they encountered in the process. This fact was reported by Binet and Henri (1894) in their historic paper which has only recently been rescued from obscurity and made available to non-French readers in an annotated partial translation by Thieman and Brewer (Note 1). While they also reported a study of memory for lists of unrelated words, their major interest was in memory for brief prose passages. The study tested the memory of 510 French grade school children, and, perhaps because of the youth of their subjects, the passages ranged in length from only 11 to 88 words. The passages were read to the subjects and recalled immediately.

They found, unsurprisingly, that the number of words recalled was a positive function of age and a negatively accelerated function of passage length. They also noted a progressive shift in error type from the predominance of synonym substitutions for shorter passages to the predominance of omission errors for longer passages. It is their analysis of the synonym substitutions, however, which is of greatest interest. They hypothesized

that synonym errors marked the loss of "verbal memory," but the retention of memory for ideas. Omission errors, on the other hand, signaled the loss of both types of memory. Since the former type of error predominated for shorter passages, their result shows that memory for words is lost before memory for ideas.

They also examined the synonym errors which occurred and noted that in the absence of verbal memory, the children tended to recall the passage in language characteristic of French school children. Thus, in addition to synonym errors, there was a tendency to simplify the syntax of the passages. Both synonym and syntactic errors were seen as instances of "l'assimilation verbologique," which is the tendency of a person to recall in language typical of that person. Early studies by Sharp (1899) and Henderson (1903) replicated and extended Binet and Henri's results.

One of the earliest demonstrations of the fact that people generally remember the meaning conveyed in a verbal message and not the exact wording of the message was provided by Buehler (1908, cited in Humphrey, 1951) using lists of unrelated sentences. Buehler saw this fact as an illustration of the more general principle that thoughts are independent of and do not follow the same processes as their sensory accompaniments. The sentence list he used was a list of 20 unrelated proverbs. When a second list of proverbs was read, some of which were similar in meaning at the metaphorical level to the proverbs in the first list ("jokers" or distractors were also included), the subjects were able to identify them with great certainty. For example, subjects who heard "When the calf is stolen, the farmer repairs the stall" in the original list would indicate it as the

appropriately related sentence when they heard "One looks to the cask when the wine escapes into the cellar" in the second list. Despite their good performance on recognizing similar meanings, however, subjects had difficulty when asked to recall the items, often producing changes in the wording. As Brewer (1974) points out, the study confounded a recognition-recall difference with its idea-wording comparison.

More recent demonstrations, however, have avoided this difficulty. English, Welborn, and Killian (1934) compared substance and verbatim memory by using different types of true-false recognition items. Substance memory was tested by summary statements or paraphrased topic sentences, while verbatim memory was tested by sentences lifted verbatim from the passage. They found that for immediate testing or a one day retention interval, verbatim memory was equal to or greater than substance memory. At longer retention intervals (4-14 weeks), however, verbatim memory declined while substance memory did not; so that, for the longer intervals, substance memory exceeded verbatim memory. Of course, this study has several weaknesses, primarily that they confounded their verbatim-substance manipulation with sentence type: summary or topic sentences versus non-summary sentences. A test only control group might also have been used to ensure that answering the questions required reading the passage.

Recently, Sachs (1967) demonstrated the superiority of substance memory using a recognition test. In her experiment, subjects listened to 24 taped passages, during each of which the tape was periodically stopped for a sentence recognition test so that the interval between the target sentence in the passage and the related test item was 0,80, or 160 syllables.

Recognition items were either semantically changed, syntactically changed (so as to leave the meaning essentially unchanged), or unchanged. It was found that with immediate testing, recognition was accurate for all types of items. After 80 or 160 syllables, however, while subjects still exceeded chance in their ability to detect syntactic changes, their performance on these items was far below the level of detection for semantic changes.

Much of the recent evidence has come from studies employing sentence lists. Fillenbaum (1966) ran a study to test memory for the idea of a sentence using antonym pairs. He used sentences in each of which was one member of an antonym pair, which was either negated or non-negated. Memory was tested using a four alternative forced choice recognition test in which the four sentences of each set were presented (2 antonyms x negated or non-negated). It was found that when errors occurred, meaning was preserved more often than changed, even though, for each item, 2 of the 3 distractors changed the meaning. This result could only obtain if memory for the meaning of the sentence remained even when memory for the words was lost.

Brewer and his colleagues have conducted a series of studies showing the superiority of memory for ideas over memory for words for both recognition and recall. Bock and Brewer (1974) used sentence lists containing sentences which were drawn from pairs of sentences which were essentially synonymous, but differed in the application of an optional syntactical transformation. They had these sentences rated for stylistic preference by an independent group of subjects. When recall was tested, errors tended to preserve meaning and to shift toward the preferred form, for all six

types of transformations employed. For only three did recall errors tend toward the untransformed sentence as a kernal plus tag theory would predict. For recognition, stylistic preference had no effect, thus adding stylistic preference in sentences to the growing list of variables which affect recall and recognition differently (others are listed by Adams, 1968; Adams & Bray, 1970; and Kintsch, 1970a, 1970b). Recognition performance was much poorer for sentences from the transformational pairs, where the item and its distractor had the same meaning but different syntax, than for control sentences which had semantically changed distractors.

Brewer and Shedletsky (Note 2) ran a recall study using sentences with parallel conceptual structures (Langacker, 1973) because these have a large number of alternative surface forms. They found that while the subjects could recall the meaning of 41% of such sentences, they only recalled the correct surface form for 15% of the sentences. Brewer (Note 3) ran a series of recall studies using sentence pairs constructed by using both items of a highly synonymous word pair. Sentences were rated for stylistic preference and a trend toward the preferred form was found in recall (Exp. II). This result obtained for both abstract and concrete synonyms (Exp. III). Thus, the study lent further support to the demonstration by Johnson, Bransford, Nyberg, and Cleary (1972) that memory for gist is as prominent for abstract as for concrete verbal material, and to their analysis that Begg and Paivio's (1969) contrary result was due to a confounding of comprehensibility with concreteness.

Another manifestation of the difference between memory for ideas and memory for words is the fact that it is much more difficult to memorize a

passage verbatim than it is to learn the substance of the passage. Several studies have compared learning rates for these two types of learning. Jones and English (1926) had subjects learn a 91 word passage which they had divided into 31 idea groups. After one reading, the subjects recalled the substance of an average of 76% of the idea groups. Verbatim memorization, however, required an average of 5.3 more trials, with some subjects failing to achieve it.

Cofer (1941) used passages of 25, 50, 100, and 150 words in a study which compared acquisition rates for substance and verbatim learning in terms of time and trials to criterion. For both time and trials, verbatim learning took longer than substance learning. Further, both time and trials to criterion increased more rapidly as a function of passage length for verbatim learning than for substance learning.

Constructive memory processes. In looking at errors which occur during recall of prose, it is clear that some errors of commission are not simply synonym substitutions, but rather result from the subject's inferring or imparting information which was not present in the original passage. These errors will be called constructive errors in this paper.

While Bartlett (1932) has certainly been the most renowned exponent of constructive memory processes (a distinction he must now, perhaps, share with Neiser, 1967), Binet and Henri (1894) once again provide one of the earliest reports of the phenomenon. They noted what they called "errors through imagination" (pp. 55-58) which occurred in the passages of more than 20 words. These errors were said to be characterized not so much by change in meaning as by addition to the meaning. These errors added details

to the original version. Thus, upon hearing "Thursday" a child would recall "last Thursday" or "Thursday evening." Likewise, "one of them" became "the youngest one", and "his parents' home" became "his home." Clearly, such errors are examples of the type of constructive errors made famous by Bartlett.

Bartlett's (1932) Remembering is a book which is a landmark in the history of cognitive psychology. The current Bartlett revival has been spurred on by the preaching of Neiser (1967) and the clever experimentation of Bransford, Barclay, Franks, McCarrell, and all the others, and it is long overdue. Although the book includes several different experimental methods and materials, it is the repeated reproduction of prose material which is relevant here. Subjects read a story twice, and repeated free recalls were collected; the first 15 minutes after reading the story, with later recalls being collected, as the opportunity arose, over retention intervals which spanned days or years. Bartlett noted that for such recalls "accuracy of reproduction, in a literal sense, is the rare exception, and not the rule" (p. 93). Most recalls showed a "tendency to abbreviate the story as a whole and also the details that are reported" (p. 72). Bartlett also, however, reported another aspect of the recalls, the constructive aspect. Some details or incidents were elaborated, with a marked dramatic effect. Although some subjects recalled only one or two striking details, others used the general setting and affective aspects of the story, along with prominent details, to construct an elaborate, if inaccurate story. A tendency by some subjects to "rationalize" the story, by filling gaps or

distorting events so as to have causal relations between events or otherwise make sense (to the reader) was also found.

While Bartlett's reconstructive notion is currently quite popular, his results and characterization of them have not gone unchallenged. Gomulicki (1956) investigated the immediate recall of prose passages from 15 to 200 words. He found that although subjects were able to recall the shorter passages verbatim, they were only able to recall the more important aspects of the longer passages. Further, he found that omissions were by far the most common error type, and he concluded that memory for passages was more accurately described as abstractive than as constructive. In fact, when judges were given both recall protocols and deliberately written abstracts of the same passages, they were little better than chance at distinguishing between them.

Gauld and Stevenson (1967) suspected that the changes and distortions in recall which Bartlett found were due to deliberate inventions by the subjects who want to "fill up gaps in their memory," a process which is separable from memory itself, rather than being a result of constructive memory processes. They ran several studies using the "War of the Ghosts" which showed that if subjects were told to recall only what was in the story and to leave gaps rather than invent if they had forgotten, this sharply reduced the number of meaning changing intrusions or additions. If the subjects were simply told to be accurate, the effect was the same as the longer injunction. Of course, these results are amenable to other interpretations. The former instructions might set up demand characteristics (Orne, 1973) which favor the production of gaps. The latter (as well as the

former) might cause the subject to raise his subjective criterion for the acceptable confidence level for response emission (Adams & Bray, 1970), causing the subject to suppress information he remembers.

While earlier studies generally failed to provide any strong support for the constructive view of memory, several recent studies have provided striking examples of reconstructive processes in action. The first in the line of studies by Bransford and his colleagues was Bransford, Barclay, and Franks' (1972) famous "turtle" study which used unrelated sentences. In it, they found that subjects who read sentences like,

Three turtles rested on a floating log, and a fish swam beneath them.

had higher algebraic recognition scores ("old," (+1) x confidence; "new," (-1) x confidence) for foils like,

Three turtles rested on a floating log, and a fish swam beneath it.

than they did for sentences where the spatial relationships was changed ("beside" substituted for "on" in both sentence and foil).

Since that time the Bransford group has conducted several studies using brief prose passages which make the same point. A study which Sue Solomon ran (cf. Bransford & McCarrell, 1974) extended the Bransford, Barclay, and Franks result to brief prose passages. Subjects heard a brief passage describing the relative locations of a pond, forest, etc. in relation to a farmyard. Before hearing the passage, they saw a picture of a farmhouse and a hill. Depending on the condition the subject was in, he saw a picture with the farmhouse on the hill or beside the hill. Subjects who saw the former picture were very likely to falsely recognize a statement which gave the location of something relative to the hill which had been

stated relative to the farmhouse (or vice versa), while subjects who saw the latter picture were not.

Johnson, Bransford, and Solomon (1973) presented subjects with brief passages such as: "John was trying to fix the bird house. He was pounding the nail when his father came out to watch him and to help him do the work." Or "It was late at night when the phone rang and a voice gave a frantic cry. The spy threw the secret document into the flames just in time, since 30 seconds longer would have been too late." They found that subjects who heard such passages were very likely to falsely recognize statements which included an unstated instrument (e.g. "John was using the hammer to fix the birdhouse. . . .") or consequence ("The spy burned the document just in time. . . ."). Subjects who heard control passages where minimal word changes altered the meaning (e.g. "looking for the nail" instead of "pounding the nail"; "pulled the secret document from the fire" instead of "threw the secret document into the fire") were much less likely to make such errors.

Brewer (Note 4) has initiated a series of studies on memory for the pragmatic implications of sentences. Pragmatic implications differ from logical implications in that while logical implications are necessarily implied under the logic of some system, pragmatic implications are not dictated by any formal system. Pragmatic implications, rather, express expectations derived from the subject's knowledge of the world. Pragmatic implications were tested for by application of the "but-not" test derived from Lakoff's (1971) analysis of the use of "but" as a denial of expectation. This test excluded sentences which had logical implications or lacked an

implicational relationship. Brewer found that pragmatic recalls constituted an average of 20% to 30% of the total recalls across 46 items and were as high as 80%-88% for some items. Examples of the most common implicational errors are:

The hungry python caught the mouse.

tended to be recalled as

The hungry python ate the mouse.

The safe cracker put the match to the fuse.

tended to be recalled as

The safe cracker lit the fuse.

Other examples of memory for pragmatic implication are provided by Harris (1974) and by Brewer and Lichtenstein (1974). Harris found that sentences like

Miss America said that she played the tuba.

are often recalled as

Miss America played the tuba.

Brewer and Lichtenstein found that for continuous antonym pairs, where negating one does not logically imply the other (unlike dichotomous antonyms, where such implication holds), subjects often recall the pragmatic implication. Thus, subjects who heard a sentence stating that something is "not hot" were apt to recall a pragmatically implied sentence stating it was "cold."

A recent wave of studies have manipulated the operation of constructive processes in discourse memory. Fredericksen (1972, 1975b) affected the occurence of constructive errors in discourse recall by instructing subjects

to remember a passage or to try to solve sociological problems posed by it. Spiro (Note 5) has demonstrated dramatic effects of telling subjects that they were in a memory experiment, or that the experiment was concerned with their reaction to interpersonal relations described in the passage. Dooling (Note 6; Sulin and Dooling, 1974) has shown that at long retention intervals, constructive errors occur for the recall of a passage about a famous person if the passage is given the appropriate title (Adolf Hitler) but not if it is given a fictitious main character (Gerald Martin).

Differences between Sentence Lists and Connected Discourse

As we have seen, it has been demonstrated with both sentence lists and connected discourse that what is remembered is usually not the exact wording of the material, but rather, its meaning. Further, memory often goes beyond the mere reproduction of meaning and exhibits constructive properties which result in material being recalled or recognized which was not present in the original message. Against this backdrop of important similarities we will now consider the question of whether or not important differences exist between sentence lists and discourse. First we will consider the studies which have directly compared the two materials, and then we will examine the question from a broader theoretical perspective.

Direct Experimental Comparisons

In this section we will examine studies which have directly compared sentence lists with connected discourse. Unfortunately, these studies have, for the most part, been atheoretical or primarily directed at other issues, so that the comparison usually does not go beyond a simple test for differences in amount recalled. There are, however, a few notable exceptions;

and, hopefully, useful suggestions for better motivated comparisons will be offered later in the paper.

A number of studies (e.g., Dooling & Lachman, 1971; Lackman & Dooling, 1968; Montague & Carter, 1973; Pompi & Lackman, 1967; Yuille & Paivio, 1969) have manipulated discourse structure by presenting the words in either normal or scrambled order. The studies to be reviewed here, however, will only be those which compare sentences in randomized lists to connected discourse. Such experiments are rather rare in the experimental psychology literature.

Anderson and Bower (1973) presented a preliminary report of two studies on verification latencies, one using connected discourse and the other using the same sentences in randomized lists. A more recent and complete journal report of these studies is also available (Anderson, 1974). The primary purpose of the studies was not to compare the two types of organization; the study was intended to test two rival hypotheses about the storage of sentence form (active or passive) information. It was found that for both types of materials, with immediate verification, verification was faster when the voice of the target sentence and the probe matched. When verification was delayed (2 min.), the effect only reached significance for the scrambled sentence condition, although the same trend occurred for the "story" condition.

Brent (1969) conducted a serial anticipation study where the materials used were organized at four different levels: isolated words, anomalous sentences, natural sentences, and paragraphs. List length was also varied (22, 34, 35, or 38 words; 7 or 10 sentences). It was found (Exp. I) that

errors (verbatim scoring) were inversely related to the level or organization. Further, although list length was a significant factor at all levels of organization, the two factors had a significant interaction resulting from the decreasing potency of the list length effect at higher levels of organization.

Just as Brent's study grew out of earlier serial anticipation studies (i.e., Brent, 1965, 1966, 1967), so Levin (1970) was led to comparing unrelated sentences to discourse by studies of the effect of embedding words in sentences in serial anticipation or paired-associate word learning tasks. Levin noted that the studies in which sentences had a facilitative effect (e.g., Clark & Bower, 1969; Levin & Posner, 1968) used a single sentence or related sentences, while those that failed to find a facilitative effect (e.g., Jensen & Rower, 1963, 1965; Levin & Rower, 1968) used unrelated sentences or phrases. Levin argued that the difference between the two sets of results obtained because the sentence or sentences served to unify the list in the former case, but not in the latter. He systematically varied the number of unrelated sentences in which words were embedded in order to test this hypothesis. Using the same 12 word list, subjects studied the words in either a single, long, connected sentence (a short story) or in 2, 3, 4, or 12 unrelated sentences (incorporating 6, 4, 3, 2, or 1 items per sentence, respectively). For both anticipation and serial learning, with fourth and fifth grade subjects, it was found total trials to criterion was greater when 3, 4, 5, or 12 sentences were used than when 1 or 2 sentences were used, as Levin had predicted.

In the educational psychology literature, there are a number of studies which have followed the line of investigation begun by Frase (1969a), in comparing name, attribute, and random organization of passages. Frase used a passage which described the attributes (moves, captures, values, etc.) of chessmen. For each of 6 chessmen, 8 attributes were described, always in a separate sentence, yielding 48 sentences. The sentences were organized by name (all 8 sentences describing a given chessman grouped together), by attribute (the 6 sentences describing a given attribute for all 6 chessmen grouped together),or randomly. Subjects received three study-free recall trials followed by a multiple choice test. For the recall tests, both attribute and name organization resulted in better performance than did random organization. The same trend appeared in the multiple choice test, but did not reach significance.

Of the studies which have followed in this line of research, not all have included a random order condition. Some have simply compared name and attribute organization (e.g., Frase, 1973a, 1973b; Friedman & Greitzer, 1972). Still, several have included random organization. Schultz & Devista (1972) used name, attribute, and random organization of a passage of 36 sentences in which six imaginary countries were given six attributes each. Once again, three study-free recall trials were employed, and once again name and attribute organization both produced better recall than did random ordering. Myers, Pezdek, and Coulson (1973) used a 25 sentence passage of five fictitious countries with five attributes each. In the two experiments which included random organization, it was found to produce poorer recall than either name or attribute organization across three study-recall trials

(Exp. I) and for both free and serial recall (Exp. II). Perlmutter and Royer (1973) replicated the effect in a study in which they used the same material as Myers et al. and the three study-test trials, but required subjects to recall the items by name, or by attribute, or by free recall.

In contrast to the studies reviewed above, two studies (Bruning, 1970; Carter, Note 7) have failed to find any difference between connected discourse and sentence lists. Bruning used 25 paragraphs of four sentences each on a fictitious African tribe. The study included three types of organization: regular prose, scrambled, and unrelated sentences in random order. The first two conditions used the same sentences, but different in their order. The third condition had unrelated sentences quasi-randomly selected from an almanac, and randomly ordered. All three, however, had the same five target sentences in the same positions in the passage. Only the target sentences were tested using a "fill in" test in which the sentence was provided, with a blank replacing the one word or number which was the correct response. While both prose and random organizations exceeded the irrelevant sentence condition, they did not differ significantly between themselves.

Bruning's failure to obtain a significant effect of sentence order may have been due to the method of testing employed in his study. The one word completion task he used is quite different from the free recall tests typically employed by the investigators cited above. Kissler and Lloyd (1972) investigated this possibility in a study which employed both a completion and a short answer test. They used eight paragraphs on math models for business, each of which had a superordinate sentence and four subordinate sentences. Half of the paragraphs were strongly related and the other half

were weakly related as determined by the coefficient of concordance for independent groups of subjects who ordered the sentences within the paragraphs. Three study-test trials were given with short answer essay questions "that required the subject to recall information from at least two sentences in the paragraph" followed by 16 one word completion items constituting the test phase. The eight paragraphs were presented in booklets in either normal or scrambled order, with sentence order being a between-subject factor. In the scrambled condition, the superordinate sentence was first in each paragraph, and the other sentences were randomly ordered within the paragraph. For the completion test, only the trials effect was significant. Thus, Bruning's failure to obtain a significant order effect using a completion test was replicated. For the short answer test, however, the order effect was highly significant, as was the trials effect. A surprising result was that the normal order was especially better than scrambled order for the unrelated paragraphs, as evidenced by a significant order x relatedness interaction. Order also interacted with trials, as the order effect was stronger for trials 2 and 3 than for the first trial.

Carter (1972) presented a 1,500 word passage on "Himoots," a fictitious South American tribe in normal or scrambled sentence order. The subjects were instructed either simply to read the passage, to read it and subjectively organize it (using notes, underlining, etc.), or to memorize a supplied outline of the material and relate the passage to it. The subjects were tested with a free recall test followed by a "cued recall" test which consisted of fill-in, short-answer, and true-false questions. The passage was divided into 210 idea units for the purpose of scoring the free recalls.

Substance scoring was employed. Free recall and "cued recall" results were analyzed together in a single analysis of variance which included a cueing factor. No effect of sentence order or of instructional set was detected. While the cueing effect reached significance, it did not interact with either order or instructions.

These results were not what Carter had anticipated. He had predicted a significant order effect and an order x cueing interaction due to a more pronounced order effect for free recall than for "cued recall." Carter and Carrier (Note 7) have pursued the matter further, and met with somewhat better results. They report two studies which employed a 1,067 word version of the Himoot passage which had four hierarchical levels (A=Himoots; B= Religion, Economic System, Appearance; C=3 subtopics for each B-level topic; D=2 subtopics of four sentences each for each C-level subtopic). In their first experiment, sentence order (normal or random) was orthogonally combined with superordinate sentence (levels A & B) placement (beginning or middle of the passage). The passage was presented in booklets with one sentence per page, and three successive repetitions of each sentence, or three repetitions of the entire passage. A free recall test was administered after a four minute filled retention interval and scored using substance scoring for 210 idea units. The result of primary interest for our purposes is that, while neither the sentence order nor superordinate placement effect reached significance, they did interact significantly. The interaction obtained because normal sentence order was better than scrambled order for the versions with superordinate sentences at the beginning. The main effect of order was significant for a clustering measure, as it had been in Carter's (1972) original study.

Carter and Carrier's second experiment used the same material, except that superordinate placement was not varied and the booklet pages contained nine or ten sentences. The passage was read one or three times. Both free recall and cued recall (C level cues) were tested, and the results were analyzed together. This time the sentence order effect was significant, but interactions with exposures and cueing revealed that the sentence order effect obtained with three exposures but not one, and with cued recall, not with uncued. A completion test was also given, but the only significant effect on this test was exposures.

A study by Lee (1965) employed a closely related order manipulation. He manipulated "supra-paragraph structure" by presenting the paragraphs of a 1,600 or 2,100 word passage on Naval discipline in randomized order, in normal order, or in normal order with added initial and final summary paragraphs, transitional paragraphs, and main headings. In all three conditions, the order of sentences within paragraphs was normal. In two experiments, subjects were tested with short essay questions worded and scored with emphasis on abstracting the main ideas presented in the passage, other questions designed to see how much detail could be recalled, and a two-alternative recognition test with targets which were sentences lifted verbatim from the passage, and distractors which were paraphrases of the targets. In both studies, the effect of organization was significant only for the test of main ideas, and then the effect was entirely due to the high level of structure, which had supplemental material totaling 380 words in the first experiment and 510 words in the second. No difference between normal and randomized paragraph order was detected in either study. This is perhaps

not too surprising, since the passages were selected on the basis of minimal interparagraph dependencies, so that the paragraphs would make sense in any order. Indeed, in the second study, when subjects were asked to rate how well the passage was organized or structured, the normal and scrambled versions did not differ significantly, although both were lower than the high structure version with the additional material.

A somewhat related field of the educational psychology literature is the effect of frame order or sequence in programmed instruction. Both detrimental effects of scrambling (e.g., Brown, 1970; Buckland, 1968) and no difference (e.g., Hamilton, 1964; Payne, Krathwohl, & Gordon, 1967) between scrambled orders have been found on criterion test performance. Other studies have found no difference on criterion test performance, but noted that "logically" ordered programs produced fewer errors and required less time for completion (e.g., Niedermeyer, Brown, & Selzen, 1969). This paper, however, will not attempt to cover the programmed instruction literature.

Theoretical Comparisons

From the brief review of experimental comparisons above, it is easy to see that connected discourse is usually easier to learn or remember than scrambled or unrelated lists of sentences. The lack of theoretical motivation in most of the studies, however, makes premature any conclusions that performance level on memory tasks constitutes the only difference. Even if that conclusion could be justified, it would still be desirable to explain the source of the difference: that is, it would still remain to determine what aspect or property of discourse, or subjects' processing of it, makes it more memorable (or easier to comprehend or learn) than sentences in lists.

So far, it seems, these questions have received little serious attention. The purpose of the next sections of this paper is to more closely consider these questions in the light of recent developments in experimental psychology. The comparisons to be made divide naturally into two types: differences between the materials themselves, or stimulus differences, and differences between the way subjects process the materials.

Stimulus differences. One of the major difficulties which has retarded the development of the discourse learning literature is the failure to make a clear distinction between the form and the content of the discourse. As a stimulus, discourse may be described at several levels of abstraction, and it is important not to confuse these levels. As a first approximation, three levels of description are sufficient: physical form, linguistic form, and content. Physical form is the level which describes the mode of presentation of the discourse, auditory or visual, and such physical characteristics of the stimulus as type face or location on a page for visual presentation and gender of the speaker or background noise for auditory presentation. Linguistic form is the level which describes the syntax of the sentences and the lexical form of the words used in the discourse. Finally, the content level describes the meaning of the discourse, and includes the representation of the informational content of the discourse, as, for example, in propositions, and the structure or interrelations of that content. Early accounts of discourse structure tended to focus on the form of the discourse. Thus, readability indices (Chall, 1958) were primarily based on word length (lexical form) and sentence length (syntactical form). More recent accounts of discourse structure have tended to focus on the content level (e.g.,

Crothers, 1972; Fredericksen, 1972, 1975a, 1975b; Kintsch, 1972, 1974). Trouble often results when these levels are confused. Idea units (Henderson, 1903) were developed in order to overcome some of the limitations of verbatim word counts (which are heavily dependent upon linguistic form) as a measure of content recall. However, when verbatim scoring is applied to the idea units, their original purpose is defeated and anomalous results may be obtained (e.g., King & Russel, 1966).

Form. For the purpose of comparing sentence lists to discourse, all variables of form, except sentence order, may be factored out by using the same sentences in the two types of material. Clearly, sentence order is not merely a form variable, but often affects content as well. Thus, while different sentence orders may sometimes convey essentially the same meaning or content (e.g., Crothers, 1972; Frase, 1969b), a random order will usually destroy much of the content of the discourse by transforming it into a sentence list. Further, in discourse, sentence order is often used to convey information such as a causal or temporal relationship. This can be illustrated by an example where clause order varies within a sentence: "John fainted and fell off his horse" means something quite different from "John fell off his horse and fainted" (adapted from Coleman, p. 277; Conference Discussion of Crothers, 1972). Likewise, "Tom's wife divorced him. He started drinking" implies a different temporal or causal relationship than "Tom started drinking. His wife divorced him." In sentence lists, since the sentences are unrelated, sentence order carries no information and is arbitrary and usually randomized.

Content. It is the relatedness of sentences in discourse, the fact that sentences in discourse are related to the theme or topic of the discourse, as well as to each other, that, to the first approximation, distinguishes discourse content from the individual meaning of all the sentences in a sentence list. Another way of expressing the difference between sentence lists and discourse with regard to theme relatedness is that connected discourse has a topic or theme, while unrelated sentence lists do not. The notion of the topic or theme of discourse is not new in psychology. Binet and Henri (1894) invoked a similar concept when comparing the learning of words in lists to words in discourse:

> The differences between memory for words and memory for connected discourse is suggested to be due to two causes. First, when one recites either, one calls to mind a series of images; however, with isolated words, the images are disorganized, or they don't interrelate naturally, and the child hears them without searching for a common association. With passages, however, the images are whole, one connected to the next, and perfectly organized. Second, when encountering isolated words, images are rapidly formed and each new one pushes the old one from consciousness, affecting its eventual strength. (p. 31, quoted in Thieman and Brewer, 1975)

The same analysis could be applied to unrelated sentences in lists; unless the subject can invent a scenario linking them, they must be learned as a disconnected list of unrelated items.

Bartlett (1932) stressed the importance of the theme in connected discourse, noting that,

> The form, plan, type, or scheme of a story seems in fact, for the ordinary, educated adult, to be the most dominant and persistent factor in the material (p. 83).

More recently, Pompi and Lachman (1967) echoed Bartlett almost exactly when they characterized the comprehension of connected discourse as the construction of "surrogate structures," which they said were "some combination of theme, image, scheme, abstract, or summary" (p. 143).

Levin (1970) stressed the unification provided by a common theme in order to explain the finding that for serial recall, anticipation, and paired-associate studies in which words were embedded in sentences, only those which employed a single sentence or related sentences demonstrated facilitation (e.g., Clark, & Bower, 1969; Levin & Rower, 1968).

When the passage is vague enough, as in studies by Dooling and Lachman (1971), Bransford and Johnson (1972), and Dooling and Mullet (1973), then the subjects will be unable to benefit from the theme relatedness of the passage, unless they are cued onto the theme with a title or picture which disambiguates the material. The theme may also be obscured by the use of indefinite articles, as in a recent study by deVilliers (1974). Brown (1973) pointed out that indefinite articles are used to introduce new referents into discourse, while a definite article is used to refer to a referent which has already been introduced. deVilliers manipulated the use of articles in order to make it more or less likely for subjects to treat a loosely knit passage about the adventures of a boy as a story rather

than as a set of unrelated sentences. When definite articles were used, the subjects assumed that referents were shared across sentences, and they were more likely to treat it as a story. With indefinite articles, subjects assumed that referents varied across sentences, and thus were more likely to treat the passage as a list of unrelated sentences. The manipulation was not entirely successful, but deVilliers used post-experimental questioning to determine how subjects had treated the passage and analyzed his data with subjects sorted on their answer. The subjects who treated the passage as a story recalled more than did those who treated it as a sentence list.

Clearly, then, a unifying theme is a key difference between connected discourse and sentence lists. The effect of the theme is to unify the passage and to provide a semantic context which guides and facilitates the comprehension, learning, and memory of sentences in the discourse. Pezdek and Royer (Note 8) conducted a pair of studies which investigated the effects of providing context on the recognition of meaning and wording changes for sentences. A list of concrete and abstract target sentences were presented either in isolation or in a paragraph context, with two context sentences preceding the target sentence. Pezdek and Royer were primarily interested in the effect of providing context on the detection of meaning changes in abstract sentences, and so their reported analyses focused on this predicted effect and the second order interaction (sentence type x context condition x test type) in which it was embedded. The predicted effect and interaction were significant in both experiments, although only marginally so in the first experiment. Of more general interest, however, is the first-order interaction of context and type of test. This interaction failed signifi-

cance in the first experiment, and was not reported in the second experiment; but an inspection of the graphed results (Figure 2) clearly shows a strong interaction; providing context aids the detection of meaning changes (subject-object reversal), but is actually detrimental to the detection of wording changes. Pezdek and Royer explain their results as showing that context increases the comprehension of a sentence and hence the probability that it will be stored primarily as a semantic interpretation of its meaning, rather than in its surface form as a list of words. Unfortunately, their study suffers from a design weakness which makes an alternative interpretation of the results possible. The increased detection of meaning changes (and also the decreased sensitivity to wording changes) in the context condition may have been due to information provided by the context sentences, rather than to any effect on the comprehension of the target sentence. A control group which studied the context sentences only, and then was given the recognition test, would be needed to evaluate this alternative source of the effect. Another solution to the problem would be to compare the context condition to the same context and target sentences presented as a randomized list.

The limitation of relatedness as an explanatory concept for discourse is that it says nothing about <u>how</u> the sentences or ideas and events expressed in them are related. The recent discourse structure models have been developed in order to begin to specify and analyze how discourse content is organized and structured. They represent attempts to take the necessary next step beyond the simple, formless concept of relatedness.

Dawes (1964, 1966) analyzed discourse structure in terms of set theoretical relations of identity, exclusion, inclusion, and disjunction. The analysis was used to study the direction of errors in recall and recognition, toward disjunction ("pseudodiscrimination") or away from it ("overgeneralization"). Dawes' analysis, however, was specific to the artificially simple material constructed for the studies, and does not appear to have the power to be applicable to more complex forms of discourse. Frase's (1969b) structural analysis is subject to the same criticisms.

A much more elaborate method of analysis was applied by Crothers (1972) who represented semantic content and structure equivalently as a directed graph or logical predicates. Although the analysis was fairly rigorous and quite complex, it proved to be seriously off the mark, since by Crothers' own interpretation (pp. 274,275), it failed to demonstrate even a trend for the overall theme to be recalled better than its details (superordinates versus subordinates in the representation) or for a major theme to be better recalled than a minor one (primary versus secondary subtrees). The failure to find better recall for the more important parts of discourse is an especially telling deficiency, because this is one of the most long standing and reliable facts of prose recall (e.g., Bartlett, 1932; Binet & Henri, 1894; Gomulicki, 1956; Johnson, 1970).

Meyer & McConkie (1973) used quite a simple method of discourse structure analysis. They had graduate students outline a passage, and then converted the outlines to tree structures. From these tree structures, three measures of the importance of an idea unit in the structure of the passage were developed: a hierarchy depth score, which measured how high in

the hierarchy the unit occurred; a units beneath score, which measured the number of units which were beneath the given unit in the hierarchy; and a combined hierarchy score, which combined the two above measures, equally weighted, into a single, unified measure. Significant effects upon recall were found for all three measures. Further, when significant effects of serial position and rated importance were found, these turned out to be largely due to the correlation of those factors with hierarchical importance. They also found that if a unit was recalled, then there was nearly a 70% chance that the unit which occurred immediately above it in the tree was also recalled, although, overall, recall was only about 23%. Further, combined hierarchy score was positively correlated with stability of recall across two recall trials.

While Meyer and McConkie's (1973) analysis has been successful, it is intuitive and informal. It also seems to be more useful for comparing the importance of units within a given structure than for comparing different structures. The analysis currently being developed by Kintsch should avoid such criticisms. Kintsch (1972) developed his propositional description of semantic content and organization as a proposal for semantic or lexical memory structure. Since that time, however, he has used it to describe the semantic content and structure of discourse material itself. He has not used his system to analyze discourse, but instead, starts with the propositional description or text base and derives prose from it. In this fashion, he has been able to keep propositional content fixed while varying syntactic complexity in order to study the effect of the latter on "reading" and "inference" times (Kintsch & Monk, 1972). He has also been able to sys-

tematically vary the number of propositions in discourse and study the effect of this factor on reading time and recall (Kintsch & Keenan, 1973). Most recently he varied the number of propositional arguments (Kintsch, 1975).

Although still in the early stages of development, Kintsch's system has already yielded promising results. Besides observing the expected effects of syntactic complexity, number of propositions, and number of arguments, it also replicated the results of Meyer and McConkie (1972), using Kintsch's more formal and objective propositional description. Kintsch's (Kintsch & Keenan, 1973) propositional rank is essentially equivalent to Meyer and McConkies' hierarchy depth score, and Kintsch's counting of descendant propositions is analogous to Meyer and McConkies' units beneath score. Kintsch's system seems to hold the most hope for future development of a formal system of description for discourse structure and content.

Propositional representation of discourse content also offers a promising basis for solving the vexing problem of scoring discourse recall. The methods normally used, verbatim or substance scoring of words, idea units, or sentences, are all too surface bound, that is too closely tied to the linguistic form of the discourse, to be completely satisfactory for measuring content recall. Propositional scoring would be similar in intent to idea units, but would lack the notorious arbitrariness of idea units (e.g., Levitt, 1956). Propositional representation is theoretically motivated and is capable both of objectivity and also of validation. Kintsch's work is as much an attempt to experimentally validate the representation of prose content provided by his text bases as it is an attempt to investigate the comprehension and memory of prose. By demonstrating that measures

derived using his system of representation exert systematic control over discourse learning and memory, Kintsch validates his system of representation both as a representation of text content, and as a basis for scoring discourse recall.

Before a particular propositional representation can be sufficiently validated to be generally accepted, however, problems general to any such system must be solved. First of all, there is the problem of how to specify what is acceptable as a realization of a proposition in recall. Second, there is the problem, with propositions with multiple arguments, of how to score partial recall. Finally, there is the problem that not all propositions are equally important. For example, the sentence

The red ball broke the window

is represented in Kintsch's system as

1. (Break, Ball, Window)
2. (Red, Ball)

It is clear that, for most purposes, the first proposition would be more important than the second. Thus, for most purposes of investigating prose learning or recall, recall will be measured best not by a simple proposition count, but rather by a weighted total, where each proposition's weight is determined by its importance. Importance can be determined by the representational system itself, as in the work of Meyer and McConkie (1972) or Kintsch (1974; Kintsch & Keenan, 1973), or it may be assigned by the investigator, if he knows what he wants the subjects to remember, as, for example in educational applications.

One major advantage of a clear distinction between form and content variables is that it leads to the concern that effects that are attributed to content variables should not be due to the confounding of form variables. Two obvious possible confoundings which might account for the effect of hierarchical importance on recall are serial position and the particular material involved.

Serial position is often confounded with hierarchical importance because structurally more important propositions tend to occur at the beginning of discourse. Thus, an observed structural importance effect might be due to a confounded serial position primacy effect. Both Meyer and McConkie (1973) and Kintsch (1974; Kintsch et al., 1975) have tested this possibility and have found that the structural importance effect remains when serial position is partialled or factored out.

While the confounding of serial position and hierarchical importance has concerned investigators of text recall, the confounding with materials used has not. Since the sentences of high structural importance are different from those of low importance, the possibility exists that the observed importance effects are in fact due to the greater learnability, memorability, or familiarity of the propositional content or sentence form of the high importance material. Clearly this possibility cannot be dismissed out of hand, and studies which test it are urgently needed. Perhaps the simplest test would be to present the same sentences in normal discourse and in randomized lists. If the structural importance effect were indeed due to the role of the material in the content structure of the discourse, then it should be observed only in the discourse condition and not in the

list condition. If the effect were due to the confounding of the material involved, then it would appear in both conditions.

The development of models of prose content has paralleled the development of network models of semantic memory. Indeed, Kintsch's system originated as a memory model and can be represented as a network, although more recently it is employed as a method of representing discourse content or text bases. Although much of the semantic memory work has gone into theory and research about the types of memory structures and process needed in order to account for latencies in verifying statements like "canaries are yellow" (e.g. Collins & Quillian, 1969, Ripps, Shoben, & Smith, 1973), there is evidence that semantic memory models may be moving towards applicability to prose learning and memory. Anderson and Bower's (1973) model HAM, which permits the embedding of proposition (context and fact) tree structures into higher order propositions, seems capable of being applied to discourse. Indeed, HAM might be able to explain deVilliers' (1974) result rather nicely by postulating that new nodes were constructed when indefinite articles were used, while links to existing nodes were constructed for definite articles, producing isolated propositions in the former case and interconnected propositions in the latter case. The structural model sketched by Rumelhart, Lindsay, and Norman (1972) and Rumelhart and Norman (Note 9), which is based on events and episodes also seems capable of dealing with discourse, although the lack of specified processes makes any application speculative. Rumelhart (Note 10) has confronted discourse directly by developing a text grammar for children's stories. Linguists as well, have become concerned with the development of text grammars (e.g., Petöfi & Riesser, 1973),

and Charniak (Note 11, Note 12) has explored the knowledge an artificial intelligence program would need to "comprehend" children's stories. Discourse structure will doubtless continue to receive an increasing share of attention in psychology, linguistics, and artificial intelligence.

Subject processing. As important as the stimulus differences between sentence lists and connected discourse are, confining an analysis of differences between the two types of material to stimulus differences would clearly leave the analysis incomplete and inadequate. Differences between how subjects process sentence lists and discourse are an equally important consideration. No truly explicit, comprehensive, and reasonable theory of how subjects process discourse is currently available, and none will be offered here. In view of the rather undeveloped state of the field, this is not surprising. Rather than reviewing existing process models of discourse comprehension, learning, and memory, such as they are, this paper will attempt to take current theoretical approaches which have broad applicability and have had considerable impact and apply them to the comparison of sentences in lists and in discourse. The approaches to be considered fall into three categories; the semantic-episodic distinction, levels of processing, and the debate over whether recognition involves retrieval processes.

Episodic and semantic memory. Tulving's (1972) paper which bisects memory and memory research has already enjoyed wide audience and influence. Tulving offered the dichotomy in order to relate the traditional verbal learning experiments, such as word lists and paired-associates to the newer semantic memory studies (e.g., Collins & Quillian, 1969; Schaffer & Wallace, 1970). In essence, episodic memory is the system which stores perceptible

aspects of events in an autobiographically based, spatio-temporal reference system. Semantic memory, on the other hand, stores formal, abstract knowledge of the language and the world without autobiographical reference. Tulving's distinction, then, is different from Atkinson's (e.g., Atkinson & Juola, 1974) partitioning of the long-term store into a lexical store and event knowledge store, in that Atkinson recapitulates the knowledge of the language versus knowledge of the world or analytic versus synthetic dichotomy while Tulving does not.

It would seem that Tulving's bisection might neatly separate sentence lists and discourse. Sentence lists are, after all, lists of verbal material, just like the word lists which are paradigmatic of episodic memory research. Discourse, on the other hand, seems a likely candidate for storage in semantic memory. In fact, Tulving cited the high incidence of "intrusions" in the recall of a prose passage (Howe, 1970) as evidence of the lack of retention of the surface form or perceptible stimulus properties of a passage when its meaning is stored in semantic memory.

On closer examination, however, this type of neat separation breaks down. In the first place, the same loss of surface form which Tulving finds typical of semantic memory occurs with sentences in lists (e.g., Bock & Brewer, 1974; Brewer, Note 3) as well as with discourse. Further, many of the passages used by psychologists investigating discourse learning and memory make storage in semantic memory extremely unlikely. It is just not reasonable to expect subjects to add "information" about a fictitious South American tribe or a serenade with electric guitar and balloons to their knowledge of the world, devoid of any autobiographical reference. Thus, retention

of meaning after loss of surfcce form, and even errors of inference (Binet & Henri, 1894) or pragmatic implication (Brewer, Note 4) must be possible for episodic memory also. Further, the extent to which the information in discourse, or, for that matter, sentence lists, gets stored in semantic memory will depend on the extent to which it is perceived by the subjects as bona fide knowledge of the world worthy of such storage. That Tulving's dichotomy does not neatly divide sentence lists and connected discourse should come as no surprise, since he cautioned that "the exercise of identifying various memory situations with episodic or semantic memory is neither simple nor particularly informative, since many tasks contain both episodic and semantic features" (p. 368).

Levels of processing. Another major contribution to thinking about subject processing is the levels of processing analysis of Anderson (1970, 1972) and Craik and Lockhart (1972; Craik, 1973). It is possible that the level of processing at which subjects typically process sentences in lists differs from the level of processing typical of discourse. In order to consider this possible difference, we must first examine the levels of processing notion.

The levels of processing concept seems to have grown naturally from earlier thought in both experimental and educational psychology. In reviewing the incidental learning literature, Mechanic (1962) and Postman (1964) concluded that intention (to learn) per se was not important, but the type of processing required by the task was crucial. Posner (1969) applied a levels of processing analysis to the abstraction of letter stimuli. Bobrow and Bower (1969), while they did not use the term levels of processing,

did hypothesize that "comprehension of a sentence aids its retention" and investigated incidental tasks designed to ensure "more reliable comprehension than does simply reading a sentence" (p. 458). Educational Psychology has produced a line of research designed to explore subject processing activities as determiners of learning from prose. Rothkopf coined the term "mathemagenic activities" (1970) which means activities which "give birth to learning" and is intended to emphasize that meaningful processing (comprehension) is necessary for any significant learning from prose to occur. Subject processing of text has been affected by means of aids such as instructional objectives (e.g., Rothkopf and Kaplan, 1972) and inserted questions (see Anderson and Biddle, 1975, for review). This paper will focus on the formulations of Anderson and Craik and Lockhart, but similar thinking has also been reflected in the work of others such as Frase (1970), Hyde and Jenkins (1969, 1973), Laberge and Samuels (1974), MacKay (1973), Mistler-Lachman (1972, 1974), and Triesman and Truxsworth (1974).

Craik and Lockhart's (1972) levels of processing has come out of the verbal learning tradition. They propose levels of processing primarily as an alternative to multistore models of memory (e.g., Atkinson & Schiffrin, 1968, 1971; Broadbent, 1958; Waugh & Norman, 1965). Craik and Lockhart's levels may be grouped into stages: sensory analysis, pattern recognition, and stimulus elaboration. Their levels of processing has been influenced by recent accounts of perception (Selfridge & Neisser, 1960; Sutherland, 1968; Treisman, 1964). They use the term "depth" of processing where deeper processing means more semantic or cognitive processing. They do not, however, insist that later stages of processing proceed in a fixed, hierarchical

order. In fact, they suggest that later stages of processing might better be characterized as "spread" of encoding rather than depth. Despite this mild disclaimer, however, they retain the term depth (implying hierarchy) and postulate that "trace persistence is a function of the depth of analysis, with deeper levels of analysis associated with more elaborate, longer lasting, and stronger traces" (p. 675).

Anderson (1970, 1972) draws upon both the verbal learning and educational psychology literatures in developing his levels of processing. He tentatively outlines the levels (for printed material) as orthographic encoding, phonological encoding, and semantic encoding. Semantic encoding is the end product of comprehension, which has been characterized by Anderson and Ortony (1975) as "constructing a particularized and elaborated mental representation." Anderson's main concern is with this deepest level of encoding because it is the level which characterizes a mature reader reading text, and because it is the level of encoding required for the attainment of reasonable and significant educational goals. He points out, however, that both orthographic and semantic codes may be retained in memory. While semantic encoding usually prodominates with sentences and discourse, so that the meaning and not the exact words themselves are retained, it is possible to learn the exact words of a passage of text (the Boy Scout Oath or Pledge of Allegiance) and even to learn a passage in an uncomprehended foreign language (perhaps a song lyric).

The important commonality in the thinking of Anderson, Craik and Lockhart, and the others, is that they characterize processing as hierarchical, with deeper levels of processing operating on the products of shallower

levels. Deeper levels are more cognitive or meaningful, and their products are more easily or better retained in memory.

The question, for the purpose of this paper, is whether subjects are more likely to comprehend or deeply analyze sentences when they appear in discourse than when they occur in unrelated lists. For example, subjects might be more inclined to try to comprehend discourse, and to rotely memorize sentences in lists. If the levels of processing analysis is correct, then several predictions are possible for both recall and recognition. For recall, substance scoring should yield higher scores for discourse than for sentence lists, as was typically found in the studies reviewed above. A clear prediction about the absolute level of verbatim recall is not possible, since verbatim recall may result not only from retention of the exact words, but also from retention of the meaning, if it is then re-expressed in the same words. Clearly, however, the proportion of substance recall which is also verbatim should be lower for discourse than for sentences. Discourse should produce more synonym substitutions, word order changes, and other changes that do not substantially change meaning than should sentence lists. In addition, errors of inference (Binet & Henri, 1894) should be more common for discourse than for sentences in lists. For recognition, it would be predicted that under neutral learning instructions recognition for semantic content should be greater for discourse than for sentence lists, but that the situation should be reversed for the recognition of the verbatim form of the sentence. Thus, levels of processing can predict Pezdek and Royer's (1973) result. Finally, experimental manipulations designed to ensure meaningful encoding of the sentences, such as imagery instructions or cover tasks which

require comprehension, should have less effect on discourse than on sentences in lists, since the former is more likely to result in meaningful encoding in the absence of such manipulations. Conversely, manipulations which discourage comprehension, such as having subjects overtly rehearse the material or monitor words for spelling errors, should have less effect on sentence lists than on discourse. These predictions certainly merit experimental investigation.

Craik and Lockhart's level of processing approach has another area of applicability to the comparison of sentence lists and discourse. They suggest that primary memory or "maintaining or recirculating information at one level of processing" by paying "continued attention to certain aspects of the stimulus" (1972, p. 676) be substituted for the short-term store. When items are recirculated in primary memory, they are not subjected to deeper levels of processing. Such recirculating can produce high immediate recall of the final items in a word list (the recency effect), but the lack of deep processing means that retention after the recirculating stops suffers, as evidenced by the negative recency effect Craik (1970) found on a final free recall after several lists. Craik and Watkins (1973, Exp. II) have since shown that inserting a 20-second unfilled retention interval, and thus greatly increasing the time in primary memory and the number of overt rehearsals of the last few items over immediate recall does nothing to alleviate the negative recency effect.

Since sentence lists are unrelated lists, like word lists, subjects may be expected to recirculate the last few items if they know that a short-term memory recall test is impending. Subjects, however, should be more

likely to process connected discourse at a deep level throughout, without employing a recirculating strategy near the end. deVilliers (1974) observed that subjects who treated his passage as a sentence list tended to output the last sentence or two first on recall, similar to the "dumping" strategy observed in short-term memory experiments with word lists (e.g., Waugh & Norman, 1965). Deese and Kaufman (1957) had previously noted the fact that subjects recalling discourse don't employ a dumping strategy. They reported a study which compared discourse recall with recall of a word list. For discourse, recall order was essentially perfectly correlated with input order. With word lists, however, subjects tended to recall last presented items first. Similar dumping has been found in the studies of short-term memory for proverbs by Glanzer and Razel (1974).

Since subjects employ recirculating and dumping with sentence lists, they should exhibit the marked short-term forgetting which occurs in the recency portion of the serial position curve. Subjects recalling discourse, however, should be much less affected by a filled retention interval because the deeper level of processing applied to the final items should make them more resistant to forgetting. Likewise, final free recall after a series of sentence lists should produce the negative recency effect, while deeper processing should prevent negative recency from occurring with discourse.

Retrieval processes in recognition. The final theoretical issue that will be examined in an effort to illuminate the comparison of sentence lists with discourse is the ongoing debate as to whether context or organizational effects in recognition memory constitute evidence for retrieval processes in recognition.

A widely held view in the verbal learning literature has been that recall involves both a generation, retrieval, or search process and a recognition test (based on familiarity, tagging, or matching with a standard), of which only the latter is necessary for recognition (e.g., Adams and Bray, 1970; Anderson & Bower, 1972; Kintsch, 1970a; McCormack, 1972; Murdock, 1968). By this view, the trace stored in memory is immediately or directly accessible (content addressable in computer parlance) in recognition, so no search process is necessary. The opposing view, championed chiefly by Tulving and his associates (e.g., Thompson, 1972; Tulving & Thompson, 1971, 1973; Watkins & Tulving, 1975) and Mandler (1972), is that search or retrieval processes play an important role in recognition as well as in recall. In other words, the memory trace is assumed not to be directly accessible during recognition, necessitating a search. Anderson and Bower's (1973) match process by which HAM recognizes sentences is an example of such a search.

A common research strategy in this area has been to identify an experimental variable which is believed to influence retrieval processes exclusively and then see if recognition is affected by such manipulation. If recognition is not affected, then the notion of searchless recognition receives a measure of support. Organization was originally thought to be such a variable. Cofer (1967), Bower (1968), Kintsch (1968), and Bruce and Fagan (1970) reported studies using word or nonsense syllable lists which showed the organizational manipulations which affected recall left recognition unaffected. Since that time, however, the evidence that organization and context affect recognition as well as recall has been steadily amassing (e.g., Bower, Clark, Lesgold, & Winzenz, 1969; D'Agostino, 1969; Franks & Bransford, 1974; Jacoby, 1972;

Lachman & Tuttle, 1965; Light & Carter-Soebel, 1970; Light & Schurr, 1973; Mandler, 1972; Thompson, 1972; Tulving & Thompson, 1971, 1973; Watkins & Tulving, 1975; Winograd & Carr, 1971).

Tulving and Mandler argue from the evidence that organization or context affects recognition that recognition includes a search process. Another explanation, however is also possible. The recognition effects can also be explained if it is assumed that organization or context affects encoding or storage and not retrieval. This argument is currently getting much play (e.g., Light, Kimble, & Pellegrino, 1975; Martin, 1975; McCormack, 1972; Pellegrino & Salzburg, 1975; Reder, Anderson, & Bjork, 1974) based on encoding variability (Martin, 1968) or stimulus sampling theory (Bower, 1972; Martin, 1972). Levels of processing (Anderson, 1970, 1972; Craik & Lockhart, 1972) can also be used to explain the effects of context or organization on recognition in terms of encoding differences.

With discourse, the effect of organizational variables on recognition has received very little attention compared with recall. Most of the work that has been done has come from Lachman and his associates (Dooling & Lachman, 1971; Lachman & Dooling, 1968; Pompi & Lachman, 1967). This series of studies has compared free recall to recognition using a rapidly paced sorting task (recognize, don't recognize) with single word stimuli. The organizational variable has been "syntactical" (normal discourse) order versus random word order. In the Pompi and Lachman study, distractors were of two types: high thematic relatedness (judged by an independent group of subjects as likely to appear in an alleged short story of which the experimental passage was a summary) and low thematic relatedness (items selected

in the same fashion for an unrelated story). No significant difference obtained for hit rate or false alarm rate, but a highly significant thematic relatedness x word order interaction occurred for the false alarm rate, owing to the fact that subjects who read the syntactical order made false positive errors much more frequently for the high thematic than the low thematic distractors, while relatedness had no effect on the false alarm rate for subjects who read the random order. Both Lachman and Dooling and Dooling and Lachman used unrelated randomly selected distractors. Lachman and Dooling varied training trials (2, 4, 6, or 8), but no test of whether the order effect was significant is reported. The recognition data is included along with recall data in a single analysis that produced many interactions which go unanalyzed. Inspection of the graph provided (of hit rates) however, reveals that if there was an order effect for recognition it was due to the groups receiving two or four training trials with the second of the two passages used. For the other conditions an order effect seems most unlikely. In the final study in the series, Dooling and Lachman add an intermediate level of organization, random phrase order. They use "vague and metaphorical" passages which are difficult or impossible to comprehend without a title and, orthogonally to the level of organization, reveal the title to half the subjects and deny it to the other half (as did Bransford and Johnson, 1972). Neither the order or the title main effect reached significance when hit rates or false alarm rates were analyzed. Both main effects, however, reached significance when difference (hits minus false alarms) were used. The order x title interaction never reached significance. Recall (free recall, verbatim word count) in all three studies was strongly

affected by order, with syntactical order producing better recall than random word order. In Lachman and Dooling's study the order effect for recall increased over trials. Dooling and Lachman's random phrase order produced a recall level intermediate between random word order and syntactical order.

Clearly these studies provide some evidence for the effect of discourse organization on recognition. The use of single word stimuli and a rapidly paced sorting task, however, may have served to minimize this effect. Since subjects tend to retain the meaning of the passage better than the exact words used, sentences would be more suitable to testing recognition. Pezdek and Royer (1973) used sentence stimuli to test for the detection of meaning or wording changes for sentences presented with or without context. As was noted above, however, the additional material presented in the context condition provided a confounding that can be avoided by an order manipulation.

An organizational factor that has not been investigated for recognition is hierarchical importance in discourse (Kintsch & Keenan, 1973; Meyer & McConkie, 1973). So far, research has shown that structurally more important phrases or propositions are recalled better in a free recall task, but the effect of structural importance on recognition needs to be tested in order to determine the generality of the effect. Once again, presenting the material in scrambled and discourse order would permit the unconfounding of the material itself with structural importance.

Although organizational or context effects do not necessarily support the notion of an important role for retrieval processes in recognition, there is another line of evidence which is somewhat more convincing. Reaction time studies beginning with Sternberg (1969a, 1969b) have shown that

the time to decide whether or not an item is a member of a target set depends on the size (number of members) of the target set. Although the target sets were quite small in the original Sternberg studies (1-6 one place digits), the effect of set size has proved to have surprising generality. Atkinson and Juola (1973, 1974; also Atkinson, Herman, and Wescourt, 1974) have obtained the size effect with target sets of 30 words or more. Landauer and Friedman (1968) and others (Landauer & Meyer, 1971; Meyer, 1970; Wilkins, 1971) have demonstrated the size effect for preexisting semantic categories by measuring the time it takes to decide whether a word (e.g., collie) belongs to a semantic category (e.g., dogs, animals).

Recently, Kintsch (1974) has extended the size effect to discourse memory. Using factual passages of 20, 40, 80, or 120 words, Kintsch found that the time required to make an affirmative response on a true-false or "verbatim" recognition test was an increasing linear function of passage length. Negative response latencies were not significantly affected by passage length.

It is odd that Kintsch should measure passage length by words, in view of his commitment to a propositional representation of discourse content. If the size effect proves to be reliable for discourse, it will provide another method for the comparison of alternative methods of representing and measuring discourse content. By constructing materials such that the number of units on two different measures of content are orthogonal to each other, the two measures could be compared. For example, four passages might be constructed with size specified by number of words and number of propositions:

Passage	Words	Propositions
1	w	p
2	w	4/3p
3	4/3w	p
4	4/3w	4/3p

The measure that demonstrated the larger effect would be competitively supported relative to the other measure. Initially, propositions might be compared to words or sentences. A more interesting comparison, however, would be to compare the n-ary propositions of Kintsch to the binary "propositions" of HAM (Anderson & Bower, 1973).

Conclusion

The purpose of this paper was to compare sentence lists and connected discourse in order to determine the relationship between these two types of materials. A brief review of the two literatures revealed that for both sentence lists and discourse, memory for meaning typically exceeds and outlasts memory for words or syntactical form, and that memory for both types of materials has been shown to involve constructive or inferential processes. Thus, important similarities between sentence lists and discourse do unquestionably exist. Unfortunately, the question of the existence of important differences between the two types of materials cannot be so confidently answered. A review of the existing literature which compares sentence lists and discourse primarily revealed that discourse is usually easier learned or better remembered. The lack of theoretical motivation in most of these studies limited the search for differences to comparing amount recalled. A few important exceptions, however, did emerge. Thus, Anderson (1974) found that

whether the voice of a test sentence matched its target affected verification latencies with delayed testing for sentences in lists but not in discourse. Pezdek and Royer (1973) found that supplying paragraph context for a target sentence aided detection of meaning changes while, perhaps, hindering detection of wording changes. deVilliers (1974) observed that subjects who perceived his passage as a sentence list tended to recall the last sentence or two first, while subjects who perceived it as a connected story recalled it in the order it was presented.

These studies suggest that important differences between sentence lists and discourse may indeed exist, in addition to the amount recalled. However, due to the embarrassing lack of empirical evidence on the issue, the nature and extent of such other differences remains unknown. It is hoped that this paper has helped to point the way to filling this void in our knowledge of human learning and memory. The question of what sorts of differences might be found has been examined in the broad context of stimulus differences between the two types of materials and in terms of possible differences in the way subjects process them. Suggestions of possible differences and of methods for experimentally investigating them have been offered. No detailed theoretical account of the differences between sentence lists and discourse was provided, indeed, in view of the present state of our ignorance on this matter, such an undertaking would be premature. The psychologists' careful, scientific, experimental study of human verbal learning and memory has put much time and effort into the study of nonsense syllables, word lists, and paired-associates. Only recently have sentences and discourse begun to claim their fair share of the psychologists' interest and study. While we

are amassing quite an impressive literature and body of data with both types of materials, our understanding of their relationship lags behind. It is hoped that the analysis provided in this paper, and especially the investigation of the experimental questions posed in it, will help move us closer to understanding how people comprehend, learn and remember sentences, both in lists and in discourse.

Reference Notes

1. Thiemann, T. J., & Brewer, W. F. Alfred Binet on memory for ideas. Unpublished manuscript, University of Illinois, 1975.

2. Brewer, W. F., & Shedletsky, L. J. Memory for ideas: Verb-phrase deletion and pronominalization. Unpublished manuscript, University of Illinois, 1974.

3. Brewer, W. F. Memory for ideas: Synonym substitutions. Unpublished manuscript, University of Illinois, 1974.

4. Brewer, W. F. Memory for the pragmatic implications of sentences. Unpublished manuscript, University of Illinois, 1974.

5. Spiro, R. J. Inferential reconstruction in memory for connected discourse (Tech. Rep. No. 2). Urbana, Illinois: University of Illinois, Laboratory for Cognitive Studies in Education, October 1975.

6. Dooling, D. J. Episodic and semantic aspects of memory for prose. Paper presented at the meeting of the American Psychological Association, Chicago, September 1975.

7. Carter, J. F., & Carrier, C. The role of selected organizational variables in learning from written instruction (Final Report for Project No. N00953-74-M-C044). Syracuse, N. Y.: Syracuse University, March 1974.

8. Pezdek, K., & Royer, J. M. The role of comprehension in learning concrete and abstract sentences. Unpublished manuscript, University of Massachusetts, 1973.

9. Rumelhart, D. E., & Norman, D. A. Active semantic networks as a model of human memory. Paper presented at Third Annual International Joint Conference on Artificial Intelligence. Stanford University, Palo Alto, California, May 1973.

10. Rumelhart, D. E. Notes on a schema for stories. First draft of a paper for presentation at the Carbonell Memorial Conference, Pajaro Dunes, California, May 1974.

11. Charniak, E. Toward a model of children's story comprehension.(Report AI TR-266). Cambridge, Mass.: Massachusetts Institute of Technology, Artificial Intelligence Laboratory, December 1972.

12. Charniak, E. Jack and Janet in search of a theory. Paper presented at Third Annual International Joint Conference on Artificial Intelligence. Stanford University, Palo Alto, California, May 1973.

References

Adams, J. A. Human memory. New York: McGraw-Hill, 1967.

Adams, J. A., & Bray, N. W. A closed-loop theory of paired-associate verbal learning. Psychological Review, 1970, 77, 385-405.

Anderson, J. R. Verbatim and propositional representation of sentences in immediate and long term memory. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 149-162.

Anderson, J. R., & Bower, G. H. Recognition and retrieval processes in free recall. Psychological Review, 1972, 79, 97-123.

Anderson, J. R., & Bower, G. H. Human associative memory. New York: John Wiley and Sons, 1973.

Anderson, R. C. Control of student mediating processes during verbal learning and instruction. Review of Educational Research, 1970, 40, 349-369.

Anderson, R. C. How to construct achievement tests to assess comprehension. Review of Educational Research, 1972, 42, 145-170.

Anderson, R. C., & Biddle, W. B. On asking people questions about what they are reading. To appear in G. Bower (Ed.), Psychology of learning and motivation, Vol. 9. New York: Academic Press, in press.

Anderson, R. C., & Ortony, A. On putting apples into bottles: A problem of polysemy. Cognitive Psychology, 1975, 7, 167-180.

Atkinson, R. C., Herrman, D. J., & Wescourt, K. T. Search processes in recognition memory. In R. L. Solso (Ed.), Theories in cognitive psychology: The Loyola Symposium. Potomac, Maryland: Lawrence Erlbaum Associates, 1974.

Atkinson, R. C., & Juola, J. F. Factors influencing speed and accuracy in word recognition. In S. Kornblum (Ed.), Attention and performance (Vol. 4). New York: Academic Press, 1973.

Atkinson, R. C., & Juola, J. F. Search and decision processes in recognition memory. In R. C. Atkinson, D. Krantz, R. D. Luce, and P. Suppes (Eds.), Contemporary developments in mathematical psychology. San Francisco: Freeman, 1974.

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds.), The psychology of learning and motivation: Advances in research and theory (Vol. 2). New York: Academic Press, 1968.

Atkinson, R. C., & Shiffrin, R. M. The control of short-term memory. Scientific American, 1971, 224, 82-89.

Bartlett, F. C. Remembering: A study in experimental and social psychology. London: C.U.P., 1932.

Begg, I. R., & Paivio, A. Concreteness and imagery in sentence meaning. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 821-827.

Binet, A., & Henri, V. La mémoire des phrases (Mémoire des idées). L'Anée Psychologique, 1894, 1, 24-59.

Bobrow, S. A., & Bower, G. H. Comprehension and recall of sentences. Journal of Experimental Psychology, 1969, 80, 455-461.

Bock, J. K., & Brewer, W. F. Reconstructive recall in sentences with alternative surface structures. Journal of Experimental Psychology, 1974, 103, 837-843.

Bower, G. H. Stimulus-sampling theory of encoding variability. In A. W. Melton and E. Martin (Eds.), Coding processes in memory. Washington: V. H. Winston and Sons, 1972.

Bower, G. H., Clark, M. C., Lesgold, A. M., & Winzenz, D. Hierarchical retrieval schemes in recall of categorized word lists. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 323,343.

Bransford, J. D., Barclay, J. R., & Franks, J. J. Semantic Memory: A constructive versus interpretive approach. Cognitive Psychology, 1972, 3, 193-209.

Bransford, J. D., & Johnson, M. K. Contextual prerequisites for understanding: Some investigations of comprehension and recall. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 717-726.

Bransford, J. D., & McCarrell, N. S. A sketch of a cognitive approach to comprehension: Some thoughts about understanding what it means to comprehend. In W. B. Weimer and D. S. Palermo (Eds.), Cognition and the symbolic processes. Potomac, Maryland: Erlbaum, 1974.

Bray, N. W., & Batchelder, W. H. Effects of instructions and retention interval on memory of presentation mode. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 367-374.

Brent, S. B. Organizational factors in learning and remembering: Functional unity of the interpolated task as a factor in retroactive interference. American Journal of Psychology, 1965, 78, 403-413.

Brent, S. B. Linguistic and nonlinguistic processes in learning and memory-interference. American Journal of Psychology, 1966, 79, 181-194.

Brent, S. B. Linguistic unity, list length, and rate of presentation in serial anticipation learning. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 70-79.

Brewer, W. F. The problem of meaning and interrelations of the higher mental processes. In W. B. Weimer and D. S. Palermo (Eds.), Cognition and the symbolic processes. Potomac, Maryland: Erlbaum, 1974.

Brewer, W. F., & Lichtenstein, E. H. Memory for marked semantic features versus memory for meaning. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 172-180.

Broadbent, D. E. Perception and communication. New York: Pergamon Press, 1958.

Brown, J. L. Effects of logical and scrambled sequences in mathematical materials on learning with programmed instructional materials. Journal of Educational Psychology, 1970, 61, 41-45.

Brown, R. W. A first language: The early stages. Cambridge, Mass.: Harvard University Press, 1973.

Bruce, D., & Fagan, R. L. More on the recognition and free recall of organized lists. Journal of Experimental Psychology, 1970, 85, 153-154.

Bruning, R. H. Short-term retention of specific factual information in prose contexts of varying organization and relevance. Journal of Educational Psychology, 1970, 61, 186-192.

Buckland, B. P. R. The ordering of frames in a linear program. Programmed Learning and Educational Technology, 1968, 5, 197-205.

Bühler, K. Tatsachen und Probleme zu einer Psychologie der Denkvorgänge. III. Über Gedankenerinnerrengen. Archive für die Gesamte Psychologie, 1908, 12, 24-92.

Carter, J. F. The role of organization in the recall of facts from prose. Unpublished doctoral dissertation. University of Illinois, 1972.

Chall, J. S. Readability: An appraisal of research and application. Columbus: Bureau of Education Research, Ohio State University, 1958.

Chomsky, N. Syntactic structures. The Hague: Mouton, 1957.

Clark, M. C., & Bower, G. H. Narrative stories as mediators for serial learning. Psychonomic Science, 1969, 14, 181-182.

Cofer, C. N. A comparison of logical and verbatim learning of prose passages of different lengths. American Journal of Psychology, 1941, 54, 1-20.

Cofer, C. N. Does conceptual organization influence the amount retained in immediate free recall? In B. J. Kleinmuntz (Ed.), Concepts and the structure of memory. New York: Wiley, 1967.

Collins, A. M., & Quillian, M. R. How to make a language user. In E. Tulving and W. Donaldson (Eds.), Organization of memory. New York: Academic Press, 1972.

Craik, F. I. M. The fate of primary memory items in free recall. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 143-148.

Craik, F. I. M. A "levels of analysis" view of memory. In P. Pliner, L. Krames, and T. Alloway (Eds.), Communication and affect: Language and thought. New York: Academic Press, 1973.

Craik, F. I. M., & Lockhart, R. S. Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 671-684.

Craik, F. I. M., & Watkins, M. J. The role of rehearsal in short-term memory. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 599-608.

Crothers, E. J. Memory structure and the recall of discourse. In J. B. Carroll and R. O. Freedle (Eds.), Language comprehension and the acquisition of knowledge. Washington, D. C.: Winston, 1972.

D'Agostino, P. R. The blocked-random effect in recognition and free recall of organized lists. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 815-820.

Dawes, R. M. Cognitive distortion. Psychological Reports (Monograph Supplement 4-VI4), 1964, 14, 443-459.

Dawes, R. M. Memory and the distortion of meaningful written material. British Journal of Psychology, 1966, 57, 77-86.

Deese, J. From the isolated verbal unit to connected discourse. In C. N. Cofer (Ed.), Verbal learning and verbal behavior. New York: McGraw-Hill, 1961.

Deese, J., & Kaufman, R. A. Serial effects in recall of unorganized and sequentially organized verbal material. Journal of Experimental Psychology, 1957, 54, 180-187.

deVilliers, P. A. Imagery and theme in recall of connected discourse. Journal of Experimental Psychology, 1974, 103, 263-268.

Dooling, D. J., & Lachman, R. Effects of comprehension on the retention of prose. Journal of Experimental Psychology, 1971, 88, 216-222.

Dooling, D. J., & Mullet, R. L. Locus of thematic effects in retention of prose. Journal of Experimental Psychology, 1973, 97, 404-406.

Ebbinghaus, H. Über das gedächtnis. Leipzig: Duneker, 1885. Translation by H. Ruges and C. E. Bussenius, Memory. New York: Teachers College, Columbia University, 1913.

English, H. B., Webborn, E. L., & Killian, C. D. Studies in substance memorization. Journal of General Psychology, 1934, 11, 233-260.

Fillenbaum, S. Memory for gist: Some relevant variables. Language and Speech, 1966, 9, 217-227.

Franks, J. J., & Bransford, J. D. Memory for syntactic form as a function of semantic context. Journal of Experimental Psychology, 1974, 103, 1037-1039.

Frase, L. T. Paragraph organization of written materials: The influence of conceptual clustering upon the level and organization of recall. Journal of Educational Psychology, 1969, 60, 394-401. (a)

Frase, L. T. Structural analysis of the knowledge that results from thinking about text. Journal of Educational Psychology, 1969, 60, No. 6, Part 2. (b)

Frase, L. T. Boundary conditions for mathemagenic activities. Review of Educational Research, 1970, 40, 337-347.

Frase, L. T. Integration of written text. Journal of Educational Psychology, 1973, 65, 252-261. (a)

Frase, L. T. Sampling and response requirements of adjunct questions. Journal of Educational Psychology, 1973, 65, 2, 273-278. (b)

Fredericksen, C. H. Effects of task-induced cognitive operations on comprehension and memory processes. In J. B. Carroll and R. O. Freedle (Eds.), Language comprehension and the acquisition of knowledge. Washington, D. C.: Winston, 1972.

Fredericksen, C. H. Acquisition of semantic information from discourse: Effects of repeated exposures. Journal of Verbal Learning and Verbal Behavior, 1974, 14, 158-169. (a)

Fredericksen, C. H. Effects of context-induced processing operations on semantic information acquired from discourse. Cognitive Psychology, 1975, 7, 139-166. (b)

Freidman, M. P., & Greitzer, F. L. Organization and study time in learning from reading. Journal of Educational Psychology, 1972, 63, 609-616.

Gauld, A., & Stephenson, G. M. Some experiments relating to Bartlett's theory of remembering. British Journal of Psychology, 1967, 58, 39-49.

Glanzer, M., & Razel, M. The size of the unit in short-term storage. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 114-131.

Gomulicki, B. R. Recall as an abstractive process. Acta Psychologica, 1956, 12, 77-94.

Hamilton, N. R. Effects of logical versus random sequencing of items in an autoinstructional program under two conditions of covert response. Journal of Educational Psychology, 1964, 55, 258-266.

Harris, R. J. Memory and comprehension of implications and inferences of complex sentences. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 217-227.

Henderson, E. N. Memory for connected trains of thought. Psychological Monographs, 1903, 5 (Whole No. 23).

Hintzman, D. L., Block, R. A., and Inskeep, N. R. Memory for mode of input. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 741-749.

Howe, M. J. A. Repeated presentation and recall of meaningful prose. Journal of Educational Psychology, 1970, 61, 214-219.

Humphrey, G. Thinking: An introduction to its experimental psychology. New York: Wiley, 1951.

Hyde, T. S., & Jenkins, J. J. Differential effects of incidental tasks on the organization of recall of a list of highly associated words. Journal of Experimental Psychology, 1969, 82, 472-481.

Hyde, T. S., & Jenkins, J. J. Recall for words as a function of semantic, graphic, and syntactic orienting tasks. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 471-480.

Jacoby, L. L. Effects of organization on recognition memory. Journal of Experimental Psychology, 1972, 92, 325-331.

Jensen, A, R., & Rower, W. D. Verbal mediation in paired-associate and serial learning. Journal of Verbal Learning and Verbal Behavior, 1963, 1, 346-352.

Jensen, A. R., & Rower, W. D., Jr. Syntactical mediation of serial and paired-associate learning as a function of age. Child Development, 1965, 36, 601-608.

Johnson, R. E. Recall of prose as a function of the structural importance of linguistic units. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 12-20.

Johnson, M. K., Bransford, J. D., Nyberg, S. E., & Cleary, J. J. Comprehension factors in interpreting memory for abstract and concrete sentences. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 451-454.

Johnson, M. K., Bransford, J. D., & Solomon, S. K. Memory for tacit implications of sentences. Journal of Experimental Psychology, 1973, 98, 203-205.

Katz, J. J., & Fodor, J. A. The structure of a semantic theory. Language, 1963, 39, 170-210.

King, D. J., & Russel, G. W. A comparison of rote and meaningful learning of connected meaningful material. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 478-483.

Kintsch, W. Recognition and free recall of organized lists. Journal of Experimental Psychology, 1968, 78, 481-487.

Kintsch, W. Models for free recall and recognition. In D. A. Norman (Ed.), Models of human memory. New York: Academic Press, 1970. (a)

Kintsch, W. Learning, memory and conceptual processes. New York: Wiley, 1970. (b)

Kintsch, W. Notes on the structure of semantic memory. In E. Tulving and W. Donaldson (Eds.), Organization of memory. New York: Academic Press, 1972.

Kintsch, W. The representation of meaning in memory. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1974.

Kintsch, W., & Keenan, J. Reading rate and retention as a function of the number of propositions in the base structure of sentences. Cognitive Psychology, 1973, 5, 257-274.

Kintsch, W., & Monk, D. Storage of complex information in memory: Some implications of the speed with which inferences can be made. Journal of Experimental Psychology, 1972, 94, 25-32.

Kissler, G. R., & Lloyd, K. E. Effect of sentence interrelation and scrambling on the recall of factual information. Journal of Educational Psychology, 1973, 64, 187-190.

Kolers, P. A., & Ostry, D. J. Time course of loss of information regarding pattern analyzing operations. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 599-612.

Laberge, D., & Samuels, S. J. Toward a theory of automatic information processing in reading. Cognitive Psychology, 1974, 6, 293-323.

Lachman, R., & Dooling, D. J. Connected discourse and random strings: Effects of number of inputs on recognition and recall. Journal of Experimental Psychology, 1968, 77, 517-522.

Lachman, R., & Tuttle, A. V. Approximation to English and short-term memory: Construction or storage? Journal of Experimental Psychology, 1965, 70, 386-393.

Lakoff, R. If's, and's, and but's about conjunction. In C. J. Fillmore and D. T. Langendoen (Eds.), Studies in linguistic semantics. New York: Holt, Rinehart, and Winston, 1971.

Landauer, T. K., & Freedman, J. L. Information retrieval from long-term memory: Category size and reaction time. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 291-295.

Landauer, T. K., & Meyer, D. E. Category size and semantic memory retrieval. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 539-549.

Langacker, R. W. Language and its structure. (2nd ed.) New York: Harcourt, Brace, Jovanovich, 1973.

Lee, W. Supra-paragraph prose structure: Its specification, perception, and effects on learning. Psychological Reports, 1965, 17, 135-144.

Levin, J. R. Verbal organizations and the facilitation of serial learning. Journal of Educational Psychology, 1970, 61, 110-117.

Levin, J. R., & Rower, W. D., Jr. Verbal organization and the facilitation of serial learning. Journal of Educational Psychology, 1968, 59, 186-190.

Levitt, E. E. A methodological study of the preparation of connected verbal stimuli for quantitative memory experiments. Journal of Experimental Psychology, 1956, 52, 33-38.

Light, L., & Carter-Soebel, L. Effects of changed semantic context on recognition memory. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 1-11.

Light, L. L., Kimble, G. A., & Pellegrino, J. W. Comments on episodic memory: When recognition fails, by Watkins and Tulving. Journal of Experimental Psychology: General, 1975, 104, 30-36.

Light, L. L., & Schurr, S. C. Context effects in recognition memory: Item order and unitization. Journal of Experimental Psychology, 1973, 100, 135-140.

Light, L. L., Stansbury, C., Rubin, C., & Linde, S. Memory for modality of presentation: Within-modality discrimination. Memory and Cognition, 1973, 1, 395-400.

MacKay, D. G. Aspects of the theory of comprehension, memory and attention. Quarterly Journal of Experimental Psychology, 1973, 25, 22-40.

Mandler, G. Organization and recognition. In E. Tulving and W. Donaldson (Eds.), Organization of memory. New York: Academic Press, 1972.

Martin, E. Stimulus meaningfulness and paired-associate transfer: An encoding variability hypothesis. Psychological Review, 1968, 75, 421-441.

Martin, E. Stimulus encoding in learning and transfer. In A. W. Melton and E. Martin (Eds.), Coding processes in human memory. Washington, D. C.: Winston, 1972.

Martin, E. Generation-recognition retrieval theory and the encoding specificity principle. Psychological Review, 1975, 82, 150-153.

McCormack, P. D. Recognition memory: How complex a retrieval system? Canadian Journal of Psychology/Review of Canadian Psychology, 1972, 26, 19-41.

Mechanic, A. Effect of orienting task, practice, and incentive on simultaneous incidental and intentional learning. Journal of Experimental Psychology, 1962, 64, 393-399.

Mehler, J. S. Some effects of grammatical transformations on the recall of simple sentences. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 346-351.

Meyer, B. J. F., & McConkie, G. W. What is recalled after hearing a passage? Journal of Educational Psychology, 1973, 65, 109-117.

Meyer, D. E. On the representation and retrieval of stored semantic information. Cognitive Psychology, 1970, 1, 242-300.

Miller, G. A. Some psychological studies of grammar. American Psychologist, 1962, 17, 748-762.

Miller, G. A., & Selfridge, J. A. Verbal context and the recall of meaningful material. American Journal of Psychology, 1950, 63, 176-186.

Mistler-Lachman, J. L. Levels of comprehension in processing of normal and ambiguous sentences. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 614-623.

Mistler-Lachman, J. L. Comprehension depth and sentence memory. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 98-106.

Montague, W. E. Elaborative strategies in verbal learning and memory. In G. H. Bower (Ed.), The psychology of learning and motivation (Vol. 6). New York: Academic Press, 1972.

Montague, W. E., & Carter, J. F. Vividness of imagery in recalling connected discourse. Journal of Educational Psychology, 1973, 64, 72-75.

Murdock, B. B., Jr. Modality effects in short-term memory: Storage or retrieval? Journal of Experimental Psychology, 1968, 77, 79-86.

Myers, J. L., Pezdek, K., & Coulson, D. Effect of prose organization upon free recall. Journal of Educational Psychology, 1973, 65, 313-320.

Neisser, U. Cognitive psychology. New York: Appleton-Century-Crofts, 1967.

Niedermeyer, F. C. The relevance of frame sequence in programmed instruction: An addition to the dialogue. AV Communication Review, 1968, 16, 301-317.

Orne, M. T. Communication by the total experimental situation: Why it is important, how it is evaluated, and its significance for the ecological validity of findings. In P. Pliner, L. Krames, and T. Alloway (Eds.), Communication and affect: Language and thought. New York: Academic Press, 1973.

Payne, D. A., Krathwohl, D. R., & Gordon, J. The effect of sequence on programmed instruction. American Educational Research Journal, 1967, 4, 125-132.

Pellegrino, J. W., & Salzberg, P. M. Encoding specificity in cued recall and context recognition. Journal of Experimental Psychology: Human Learning and Memory, 1975, 1, 538-548.

Perlmutter, J., & Royer, J. M. Organization of prose materials: Stimulus, storage, and retrieval. Canadian Journal of Psychology, 1973, 27, 200-209.

Petöfi, P. J. S., & Reiser, H. (Eds.). Studies in text grammar. Dondrecht, Holland: D. Reidel Publishing Company, 1973.

Pompi, K. F., & Lachman, R. Surrogate processes in the short-term retention of connected discourse. Journal of Experimental Psychology, 1967, 75, 143-150.

Posner, M. I. Abstraction and the process of recognition. In G. H. Bower and J. T. Spence (Eds.), The psychology of learning and motivation: Advances in research and theory (Vol. 3). New York: McGraw-Hill, 1969.

Postman, L. Short-term memory and incidental learning. In A. W. Melton (Ed.) Categories of human learning. New York: Academic Press, 1964.

Reder, L. M., Anderson, J. R., & Bjork, R. A. A semantic interpretation of encoding specificity. Journal of Experimental Psychology, 1974, 102, 648-656.

Ripps, L. J., Shoben, E. J., & Smith, E. E. Semantic distance and the verification of semantic relations. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 1-20.

Rothkopf, E. Z. The concept of mathemagenic activities. Review of Educational Research, 1970, 40, 325-336.

Rothkopf, E. Z, & Kaplan, R. Exploration of the effect of density and specificity of instructional objectives on learning from text. Journal of Educational Psychology, 1972, 63, 295-302.

Rumelhart, D. E., Lindsay, P. H., & Norman, D. A. A process model for long-term memory. In E. Tulving and W. Donaldson (Eds.), Organization and memory. New York: Academic Press, 1972.

Sachs, J. S. Recognition memory for syntactic and semantic aspects of connected discourse. Perception and Psychophysics, 1967, 2, 437-442.

Schultz, C. B., & DiVesta, F. J. The effects of passage organization and note taking on recall of textual materials. Journal of Educational Psychology, 1972, 63, 244-252.

Selfridge, O. G., & Neisser, U. Pattern recognition by machine. Scientific American, 1960, 203, 60-68.

Sharp, S. E. Individual psychology: A study in psychological method. American Journal of Psychology, 1899, 10 , 329-391.

Sternberg, S. The discovery of processing stages: Extensions of Donder's method. Acta Psychologica, 1969, 30, 276-315. (a)

Sternberg, S. Memory-scanning: Mental processes revealed by reaction-time experiments. American Scientist, 1969, 57, 421-457. (b)

Sutherland, N. S. Outlines of a theory of visual pattern recognition in animals and man. Proceedings of the Royal Society, Series B, 1968, 171, 297-317.

Thompson, D. M. Context effects in recognition memory. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 497-511.

Triesman, A. Monitoring and storage of irrelevant messages in selective attention. Journal of Verbal Learning and Verbal Behavior, 1964, 3, 449-459.

Triesman, A., & Truxworth, J. Immediate and delayed recall of sentences after processing at different levels. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 38-44.

Tulving, E. Episodic and semantic memory. In E. Tulving and W. Donaldson (Eds.), Organization and memory. New York: Academic Press, 1972.

Tulving, E., & Thompson, D. M. Retrieval processes in recognition memory: Effects of associative context. Journal of Experimental Psychology, 1971, 87, 110-124.

Tulving, E., & Thompson, D. M. Encoding specificity and retrieval processes in episodic memory. Psychological Review, 1973, 80, 352-373.

Tulving, E., & Watkins, M. J. Continuity between recognition and recall. American Journal of Psychology, 1973, 86, 739-748.

Watkins, M. J., & Tulving, E. Episodic memory: When recognition fails. Journal of Experimental Psychology: General, 1975, 104, 5-29.

Waugh, N. C., & Norman, D. A. Primary memory. Psychological Review, 1965, 72, 89-104.

Wilkins, A. Conjoint frequency, category size, and categorization time. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 382-385.

Winograd, E., & Conn, C. P. Evidence from recognition memory for specific encoding of unmodified homographs. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 702-706.

Yngve, V. H. A model and an hypothesis for language structure. Proceedings of the American Philosophical Society, 1960, 104, 444-466.

Yuille, J. C., & Paivio, A. Abstractness and the recall of connected discourse. Journal of Experimental Psychology, 1969, 82, 467-471.